Dynamic Matching: Centralized Early Admissions and Affirmative Actions

Tong Wang[*] Congyi Zhou[†]

April 29, 2020

Abstract

In the past decade, one of largest changes in the Chinese high school admission is adopting a Chinese version of affirmative actions. The Chinese affirmative action involves a flexible and privilege-based school choice method that has gained profound fame through time. Specifically, in the admission procedure, several designated students receive a privilege (lump-sum extra scores) apart from their exam scores. Two popular procedures are used to determine who could receive this privilege: one involves an early selection before the normal admission procedure, and the other adjusts the priority based on the rank-ordered list submitted by schools also in the normal admission procedure. However, neither of these two mechanisms is minimal responsive or respects improvement. This study proposes a new privilege-based mechanism, namely, the student optimal type-privilege mechanism (SOTPM), which is strategy-proof, minimal responsive and respects improvement. We also combine a administrative record with survey data from China to estimate students' preferences and conduct the counterfactual analysis. We find that SOTPM can considerably increase the chances of students from low quality middle schools to gain entry into good quality high schools.

[*]Tong Wang, The Waseda Institute for Advanced Study, tongw@aoni.waseda.jp

[†]Congyi Zhou, Wilf Family Department of Politics, New York University, cz536@nyu.edu

1 Introduction

The analysis of centralized school choice mechanisms has become a key focus of research in market design (Abdulkadiroglu and Sönmez 2003). In extant literature on the school choice problem, affirmative actions in the public school admission begin to receive increasing attention. Yet unlike the quota-based affirmative action, many Chinese cities have implemented a privilege-based affirmative action system, in which students from a specific schools (or groups) receive an extra scores in their entrance exams. Through this process, students from under-representative groups may have better chances of getting into good schools, but the admission is unguaranteed. In this project, we study the theoretical properties of the most popular school choice mechanisms to implement the affirmative action policy in China, and empirically test this policy using a high school admission data set.

The high school admission system in each individual Chinese city is a merit based matching process. The entrance exam score is the unique priority used as a centralized matching mechanism. In many cities, the disparity in the quality of middle schools turned certain middle schools into a feeding school for the best high schools, whereas others can only send a few students to those high schools. Given the extremely intense competition in the college entrance exam, admission into a good high school is probably a unique path for most middle school graduates to successfully attend college in the future. To offer other opportunities for students from low-quality middle schools, many cities implement their affirmative action policies in the high school admission process. The affirmative action in high school admission is unrelated to ethnic groups or household income. More precisely, each high school provides a number of privileges, allowing a set middle school specific quota. A student who receives this privilege will obtain a lump-sum of extra points in the entrance exam. The students who receive the privilege are called the *indexed* students; thus, the affirmative action is also referred to as "the indexed student" policy in China.

Two popular affirmative action mechanisms are adopted in China. First is the Chinese early selection mechanism (CESM), which involves a two-step dynamic matching procedure. In this mechanism, each middle school will use an early selection stage to determine who may receive the privilege. In general, the middle school ranks its students by their weighted average GPAs and then use a serial dictatorship process to assign the privilege to students. Then, in the normal admission stage, an adjusted Chinese parallel mechanism is adopted. Two differences emerge between this mechanism and the standard Chinese parallel mechanism (Chen and Kesten 2017). First, students who receive this privilege must choose the high school he/she wishes to attend in the early selection as the first choice in her rank-ordered list. Second, if an indexed student is rejected by her first choice after adding the extra privilege score to her exam score, then, she will lose the extra privilege points for the rest of her choices. This complicated matching procedure in the CESM has two key traits: the existence of the early selection stage and the variety of priority with matching rounds, that is, the indexed students' privilege vanishes after the first round matching in the normal admission. These two properties cause flaws. The CESM is not minimally responsive and thus may penalize the students from one specific middle school (or group) if the privilege quota to this school is increased. Moreover, the CESM does not respect merit improvement, in which a student might be penalized in the matching outcome if her relative ranking among all students is strictly improved. In reality, parents and some school principals also complain about the potential collusion among students in the matching. For example, a middle school principal may ask some top students from this school to strategically relinquish the privilege in the early selection stage and leave these opportunities to students who are on the margin of gaining admission. Through this process, this school may send additional students into top high schools.

One mechanism adopted by some local education bureaus to overcome the flaws in the CESM is the Chinese type-specific mechanism (CTSM). This mechanism has no early selection stage. Instead, a fixed number of top students from a specific middle school will automatically receive the lump-sum extra points for their priority but only first chosen school in the ROLs. However, this mechanism fails to overcome the faults. First, even the matching algorithm follows the DA mechanism, the variety of priority scores with matching rounds still renders the CTSM a non-strategy-proof mechanism. Second, the CTSM is still not minimal responsive and does not respect merit improvement.

Although the existing mechanisms to implement affirmative action in China contain flaws, the privilege-based mechanism has its advantages. (Kojima 2012) proves an impossible theory that a quota-based affirmative action mechanism cannot satisfy the minimal responsiveness because enlarging a reserved quota for a minority group may trigger a rejection chain that penalizes some minority students without benefitting others. (Hafalir et al. 2013) and (Doğan and Klaus 2018) provide partial solutions to overcome this shortcoming. The current study proposes a new mechanism, namely, the student optimal type-privilege mechanism (SOTPM), which is a simple extension of the deferred acceptance (DA) mechanism. The top students in a specific middle school can receive the lump-sum extra points; then, this privilege is valid for all their choices. The matching algorithm simply follows the DA mechanism. The SOTPM is not only strategy-proof, but the matching outcomes under this mechanism are preferred by any students to any stable matching results. The SOTPM is a privilege-based mechanism. When the privilege quota for a specific school is increased, students who receive this privilege have unguaranteed admission. In other words, unless some students are better off by attending desired schools through the increase in the privilege quota, no other students will be worse off by rejecting the original assignment. Therefore, this mechanism satisfies the minimal responsiveness and respects improvement of a student's exam score.

The theoretical properties of these mechanism motivate us to investigate real-world student behavior and welfare consequences. One difficulty with any empirical analysis of the school choice problem is estimating student preferences when only the submitted applications can be observed. The reason is that, if the adopted mechanism is *not* strategy-proof, then, students have an incentive to misreport their true preferences when submitting their rank-ordered lists. Our survey, which covered nearly half of those who graduated from middle school in 2014, aimed to uncover students' true preferences and thereby somewhat counter the problems associated with assessing those preferences in the presence of strategic behavior. A comparison of survey responses and the ROLs actually submitted indicates that students also sought to increase their chances of gaining entry by strategically maintaining sufficient gaps between their ROL choices. Given that indexed students tend to gain admission by their first choices but will lose the privilege score once rejected, the average gap between their first and second choices in the ROLs is considerably larger than that for other students.

Survey results are used to estimate student preferences over schools without considering strategic behavior in ROLs. Our estimated results indicate that a 1-unit increase in school quality (see Section 4.2 for the definition) is associated with normal high-scoring girls being willing to travel an additional distance 0.54 kilometers; the corresponding distances for medium- and low-scoring girls are 0.2 km and 0.18 km, respectively. In the same situation, high-scoring boys are willing to travel an additional 2.75 km. For indexed students, if school quality increases by 1 unit, girls are willing to travel an additional distance around 0.32 km, and boys are willing to travel an additional 0.92 km.

Using the estimated student preferences, we conduct counterfactual experiments that enable assessment of how the different matching mechanisms perform. We use the simulated matching outcomes under the SOTPM mechanism as the benchmark. When that mechanism is replaced with the CESM, student welfare is increased (on average) by 7.8%, however both high- and medium-scoring students experience a welfare loss, but the low-scoring students receive considerable welfare gain on average. When SOTPM is replaced by the CTSM, then student welfare is reduced (on average) by 2%, the high-scoring students have a welfare gain by 2%, while medium- and low- scoring students have a welfare loss by less than 2%.

Considering that introducing the indexed student policy targets at providing graduates from low-quality middle schools additional opportunities to enter good high schools, we also examine the influence of different mechanisms on diversity. If no indexed student policy is implemented, 59% of the seats of the best two high schools will be taken by graduates from the top ten middle schools under the DA mechanism. Under the SOTPM, students from top ten middle schools will take less than 40% of seats of the best two high schools, while students from median level middle schools will take 51% of seats of these two schools relative to 36% under the DA mechanism. Students from the bottom ten middle schools may take 9% seats from the best two high schools relative to 5% under the DA mechanism.

The rest of this paper proceeds as follows. In Section 2 we present school choice mechanisms with the affirmative action policy in China. Section 3 provides details on the local indexed student policy's background, after which Section 4 describes our data and analyzes students' strategic behavior in the applications. We present the empirical model and our estimates of student preferences in Section 5, and in Section 6 we conduct counterfactual experiments across mechanisms. Section 7 concludes with a summary of our findings.

2 School Choice Problem with Type-Privilege

A school choice problem with type-privilege is a tuple $G = (H, M, I, \mathbf{q}, P_I, \pi)$ with:

1. a finite set of high schools, $H = \{h_1, \ldots, h_k\}$, and a finite set of middle schools (types), $M = \{m_1, \ldots, m_r\}$;

2. a set of students $I = \cup_{m \in M} I_m$, and each middle school m has finite students $I_m = \{i_m^1, \ldots, i_m^{k_m}\}$;

3. each high school h has a quota $\mathbf{q}_h = (q_h, q^p_{hm_1}, \ldots, q^p_{hm_2})$, where q_h is h's total quota, and q^p_{hm} represents particular number of privileges assigned to middle school m's students from high school h, which is known as h's m-type privilege quota. $q^p_h = \sum_{m \in M} q^p_{hm}$ is known as high school h's total privilege quota, $q^p_m = \sum_{h \in H} q^p_{hm}$ is known as middle school m's total privilege quota and $q^p_m \leq |I_m|$;

4. a list of strict student preferences $P_I = (P_{i_1}, \ldots, P_{i_n})$ over high schools and $\emptyset$, which denotes being unassigned;

5. a list of strict school priority score profiles $\pi = (\pi_{h_1}, \ldots, \pi_{h_k})$, where π_h is school h's priority scores over I.

For any student i, $hP_i\emptyset$ means school h is acceptable for i; and R_i denote the "at least as good as" relation induced by P_i. The score profile is a function $\pi_h : \cup I_m \to \mathbb{R}$. $\pi_h(i) > \pi_h(j)$

means that student i has higher priority (score) than student j in school h. We further assume the total capacity is large than the total number of students, $|I| \leq \sum_{h \in H} q_h$.

A **matching** μ is a function $\mu : I \to H \cup \emptyset$ with $|\mu^{-1}(h)| < q_h$. A matching μ is **individually rational** if there is no student i such that $\emptyset P_i \mu(i)$. A matching μ is **non-wasteful** if there is no student-school pair (i, h) such that $hP_i\mu(i)$ and $|\mu^{-1}(h)| < q_h$. A matching μ is **fair** if there is no student-school pair (i, h) such that $hP_i\mu(i)$ and $\pi_h(i) > \pi_h(j)$ for some $j \in \mu^{-1}(h)$. A matching μ is **stable** if it is individually rational, non-wasteful, and fair. A **mechanism** φ is a systematic procedure that selects a matching for each problem. Let $\varphi(G)$ denote the matching selected by φ in problem G, and a mechanism φ is stable if $\varphi(G)$ is stable for any G and, φ is **strategy-proof** if it is a dominant strategy for each student to truthfully report her preferences.

Since, we will consider the affirmative action problem, so we call a matching μ is **weakly stable** if μ is individually rational, and if $hP_i\mu(i)$, then either (a) $|\mu^{-1}(h)| = q_h$ and $\pi(j) > \pi(i)$ for all $j \in \mu^{-1}(h)$, or (b) $i \in m$, $|\mu^{-1}(h) \cap m| = q^p_{hm}$ and $\pi_h(j) > \pi_h(i)$ for all $j \in \mu^{-1}(h) \cap m$ (**weakly fair**). The definition of weakly stability is standard, except that condition (b) describes a case that there is no potential blocking within the same middle school (type) students, not between different middle schools' students. A mechanism φ is weakly stable if $\varphi(G)$ is weakly stable for any G.

2.1 Chinese Early Selection Mechanism (CESM)

Since the high school admissions in China are merit-based procedures, which is rely on students' high school entrance exam scores, we assume all high schools share the same score profile π, and simply call it "*exam score*" and drop the subscript h hereafter.

Chinese early selection mechanism has two stages, in the first stage the middle school specific privilege is determined by an early selection procedure, in which each middle school adopts a serial dictatorship to assign high school privileges to its students; in the second stage, all students are assigned to schools by an adjusted Chinese parallel mechanism.

Early Selection (ES) Stage:

First, each student i receives her score $\pi(i)$, and each middle school ranks its students by their scores. For each middle school m, the selection runs in the following way:

• In the first step, the student who ranks the first chooses either to be an indexed student of any high school h as long as $q^p_{hm} > 0$ or to give up the opportunity to be an indexed student. In the former case, the chosen school h's m-type privilege quota is deducted by one, and this student becomes an indexed student of school h.

• In the k-th step, the student who ranks the k-th in this school, chooses either to be an indexed student of any high school h as long as h's remaining m-type specific privilege is still unfilled, or to give up the opportunity to be an index student. In the former case, the chosen school h's m-type privilege quota is deducted by one, and this student become a indexed student of school h.

The selection terminates when either all type specific privilege in middle school m are run out, or every student in this middle school has received an opportunity to make a decision.

Normal Admission (NA) Stage:

The matching algorithm in this stage follows an adjusted Chinese parallel mechanism with permanent-execution vector $\mathbf{e} = (e_1, e_2, \ldots)$. Meanwhile, every student who becomes an indexed student of any high school h in the early selection stage must choose h as her first choice in the ROL. In addition to their exam scores, the indexed students receive an extra lump-sum privilege score $\bar{s}$ for their first choices, but not for other choices in the matching algorithm. More precisely, the algorithm selects the matching outcomes as follow:

Round 1:

- At the begin of this round, every indexed student's exam score is added by a lump sum score $\bar{s}$, i.e. the new priority score of an indexed student i becomes $\pi(i) + \bar{s}$. Other students' priority scores are still π. Then each student applies to her first choice. Each high school h tentatively holds the top q_h applicants based on their priority scores in the admission pool, and reject other applicants. Furthermore, each rejected indexed student's priority score is deducted by $\bar{s}$ (the priority score is changed back to π).

  In general:

- Each rejected student i who has not yet applied to her (e_1)th-choice school applies to her next-preferred school. A student who has been rejected by all her first e_1 choices does not apply to any other schools until the next round. Each school h reviews the new applicants, along with those currently held in the admission pool, and then tentatively holds the top q_h applicants in its admission pool based on the priority scores. The other applicants are rejected.

- The round terminates whenever each student either is held in a school's pool or has been rejected by all her first e_1 choices. At this point, all tentative assignments become final. For each school h, its remaining total quotas are denoted $q_{h,2}$.

In general

Round $k > 1$

- Each student applies to her $\left(\sum_{j=1}^{k-1} e_j + 1\right)$th-choice school. Then, as in Round 1, each school j tentatively holds the top $q_{h,k}$ applicants in the admission pool (again, based on priority scores).

In general:

- Each rejected student i who has not already applied to her $\left(\sum_{j=1}^{k} e_j\right)$th-choice school applies to her next-preferred school. A student who has been rejected by all her first $\sum_{j=1}^{k} e_j$ choices does not apply to any other schools until the next round. Each school j reviews the new applicants, along with those currently held in the pool, and tentatively holds the top $q_{h,k}$ applicants in its pool based on priority scores.

The algorithm terminates when each student is admitted to a school and all the tentative assignments are final.

The CESM has two main differences with most of school choice mechanisms. First, it involves a dynamic procedure by introducing the ES stage. Second, the priority score used in the mechanism varies with matching round. Next, we will analyze the deficiencies brought by these two problems. Since the Chinese parallel mechanism is a family of mechanisms. Except a special case-the DA mechanism, other mechanisms in this family is not strategy-proof. To highlight the properties of involving the dynamic procedure and variety of the priority score, we focus on the case that the DA mechanism is adopted in the NA stage. We will show that even both the ES and NA stage adopts a strategy-proof matching procedure, the CESM is still subject to deficiencies. Hereafter, we use the CESM to represents the CESM with $\mathbf{e} = (\infty, \infty, \ldots)$. The properties of the CESM with general permanent-execution vector $\mathbf{e}$ can be found in the appendix.

First, we show the CESM may have different Nash equilibrium outcomes, in which students do not change the order of their true preferences on high schools.

Example 1. There are two high schools h_1 and h_2 and two middle schools m_1 and m_2 with $\mathbf{q}_{h_1} = \{2, 0, 1\}$ and $\mathbf{q}_{h_2} = \{1, 0, 0\}$. Students $i_1 \in m_1$ and $i_2, i_3 \in m_2$ with exam scores $\pi(i_2) > \pi(i_1) > \pi(i_3)$ and privilege scores $\pi(i_2) > \pi(i_3) + \bar{s} > \pi(i_1)$. All students have the same true preferences on high schools, i.e. $h_1 P_i h_2$.

Let's consider the first strategy profiles, students report their preference as follow:

$$
\begin{aligned}
i_1 &: \text{ ES: } \cdots | \quad \text{NA: } \{h_1 P_i h_2\}, \\
i_2 &: \text{ ES: } h_1 P_i \emptyset \cdots | \quad \text{NA: } \{h_1 | h_1 P_i h_2, \cdots\}, \\
i_3 &: \text{ ES: } h_1 P_i \emptyset \cdots | \quad \text{NA: } \{h_1 | h_1 P_i h_2, \cdots\},
\end{aligned}
$$

In this profile, i_1's middle school m_1 doesn't have privilege quota from any high school, i_1 reports her true preference $h_1 P_i h_2$ in the normal admission stage; i_2 chooses to become the indexed student of h_1 rather than be a non-indexed student in the ES stage $(h_1 P_i \emptyset)$, and her

ROL in the NA stage is $h_1 P_i h_2$ if she is an indexed student of h_1; and i_3 adopts the same strategy as i_2. This strategy profile is an Nash equilibrium, and the matching outcome of the CESM is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_2\}$$

.

Now consider another strategy profile as follow:

$$\begin{aligned} i_1 : &\quad \text{ES:} \quad \cdots | \quad \text{NA: } \{h_1 P_i h_2\}, \\ i_2 : &\quad \text{ES:} \quad \emptyset P_i h_1 \cdots | \quad \text{NA: } \{h_1 | h_1 P_i h_2, \cdots\}, \\ i_3 : &\quad \text{ES:} \quad h_1 P_i \emptyset \cdots | \quad \text{NA: } \{h_1 | h_1 P_i h_2, \cdots\}. \end{aligned}$$

The only difference between this profile and the first one is i_2 chooses to be a non-indexed student in the ES stage ($\emptyset P_i h_1$). This strategy profile is also an Nash equilibrium with the matching outcome:

$$\{i_1 - h_2; i_2 - h_1; i_3 - h_1\}$$

Example 1 indicates when students report preferences without changing the order of their true preferences in both the ES and NA stage, multi-equilibra with different matching outcomes still exist. It is because the CESM involves a dynamic procedure, in which students may strategically choose whether to become an indexed student in the early selection stage. It also implies there may not be a dominant strategy for all students to choose in the CESM, hence our analysis will focus on Nash equilibrium outcomes of the CESM. The next proposition indicates the equilibrium outcome of the CESM is weakly stable.

Proposition 1. *A Nash equilibrium outcome of the CESM is weakly stable.*

The adoption of affirmative action policy is to promote students from under-representative groups into better schools. Hence the mechanism used to implement affirmative action is expected to increase (or at least not decrease) the opportunity for under-representative students getting to good schools when the privilege quota for them is increased. This expectation is referred as the minimal responsiveness(Doğan 2016). More precisely, A matching $\widetilde{\mu}$ is **Pareto inferior to μ for school** m if (i) $\mu(i) R_i \widetilde{\mu}(i)$ for every $i \in m$ and (ii) $\mu(s) P_i \widetilde{\mu}(s)$ for at least one $s \in m$. A mechanism φ is **minimally responsive** in quota if two problem G and G' are the same except a middle school m has a higher privilege quota from a high school h in G' than in G, i.e. $q_{hm}^{p'} > q_{hm}^{p}$, then the matching outcome in G' is not Pareto inferior for the students in m to that in G. Kojima (2012) illustrates any quota-based affirmative action mechanism cannot achieve minimal responsiveness. In the next example, we show that although the CESM is a privilege-based mechanism, it is still not a minimal responsive mechanism.

Example 2. There are two high schools h_1 and h_2 and two middle schools m_1 and m_2 with $\mathbf{q}_{h_1} = \{3, 1, 0\}$ and $\mathbf{q}_{h_2} = \{1, 0, 0\}$. Students $i_1, i_2, i_3 \in m_1$ and $i_4 \in m_2$ under exam scores $\pi(i_1) > \pi(i_2) > \pi(i_4) > \pi(i_3)$ and privilege scores $\pi(i_1) > \pi(i_2) > \pi(i_3) + \bar{s} > \pi(i_4)$. All students have the same true preferences on high schools, i.e. $h_1 P_i h_2$.

Students report their preferences as follow:

$$\begin{aligned} i_1, i_2: &\quad \text{ES: } \emptyset P_i h_1 \cdots | \quad \text{NA: } \{h_1 | h_1 P_i h_2 \cdots\}, \\ i_3: &\quad \text{ES: } h_1 P_i \emptyset \cdots | \quad \text{NA: } \{h_1 | h_1 P_i h_2 \cdots\}, \\ i_4: &\quad \text{ES: } \cdots | \quad \text{NA: } \{h_1 P_i h_2\}, \end{aligned}$$

This strategy profile is an Nash equilibrium, and the matching outcome of the CESM is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_1; i_4 - h_2\}.$$

Now if m_1 receives one more privilege seat from h_1, i.e. $\mathbf{q}_{h_1} = \{3, 2, 0\}$, and students report their preferences as follow:

$$\begin{aligned} i_1, i_2: &\quad \text{ES: } h_1 P_i \emptyset \cdots | \quad \text{NA: } \{h_1 | h_1 P_i h_2 \cdots\}, \\ i_3: &\quad \text{ES: } h_1 P_i \emptyset \cdots | \quad \text{NA: } \{h_1 | h_1 P_i h_2 \cdots\}, \\ i_4: &\quad \text{ES: } \cdots | \quad \text{NA: } \{h_1 P_i h_2\}, \end{aligned}$$

This strategy profile is an Nash equilibrium, and the matching outcome of the CESM is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_2; i_4 - h_1\}.$$

i_3 is strictly worse off in the second case and no other students in m_1 are better off when h_1 gives one more privilege seat to m_1.

Example 2 not only indicates that increasing the privilege quota may hurt a specific type of students, but also shows an interesting phenomenon that the matching outcome can be affected by a group collaboration. In the first equilibrium of this example, i_1 and i_2 give the opportunity to be the indexed student of h_1 to i_3, this strategy make the students from m_1 take all three seats in h_1. However, in the second equilibrium, if both i_1 and i_2 take the privilege of h_1, then i_3 loses the chance to get into her favorite school even the privilege quota is increased for her middle school. This example implies that in order to send more students into a good school, a middle school (type) may strategically arranges its most competitive (high scoring) students give up the privileges and let the students whose scores are the margin take advantage these privileges. Due to this group collaboration, we will show in the next example that a student may be worse off in a matching outcome even her priority standing is strictly improved.

One of the most important parameters is the exam score π. We expect that a reasonable mechanism would not penalize a student when her standing in the list is improved. Given two exam score lists π_2 and π_1, π_2 is an **unambiguous improvement** for student i over π_1 if (i) the relative ranking among all students except i remains the same between π_2 and π_1, and (ii) the standing of i is strictly better under π_2 than under π_1. Then a mechanism **respects improvement** if a student never receives a strictly worse assignment as a result of an unambiguous improvement of her priority ranking. Unfortunately, the equilibrium outcome of the CESM may not be respect improvement.

Example 3. Equilibrium outcome of the CESM may not be respect improvement: There are two high schools h_1 and h_2 and two middle schools m_1 and m_2 with $\mathbf{q}_{h_1} = \{4, 2, 0\}$ and $\mathbf{q}_{h_2} = \{2, 0, 0\}$. Students $i_1, i_2, i_3, i_4 \in m_1$ and $i_5, i_6 \in m_2$ under exam scores $\pi_1(i_1) > \pi_1(i_2) > \pi_1(i_3) > \pi_1(i_4) > \pi_1(i_5) > \pi_1(i_6)$ and privilege scores $\pi_1(i_1) > \pi_1(i_2) > \pi_1(i_3) + \bar{s} > \pi_1(i_4) + \bar{s} > \pi_1(i_5) > \pi_1(i_6)$. All students have the same true preferences on high schools, i.e. $h_1 P_i h_2$.

Students report preferences as follow:

$$\begin{aligned} i_1, i_2 :&\quad \text{ES:}\quad \emptyset P_i h_1 \cdots | \quad \text{NA: } \{h_1 P_i h_2\},\\ i_3, i_4 :&\quad \text{ES:}\quad h_1 P_i \emptyset \cdots | \quad \text{NA: } \{h_1 | h_1 P_i h_2, \cdots\},\\ i_5, i_6 :&\quad \text{ES:}\quad \cdots | \quad \text{NA: } \{h_1 P_i h_2\}. \end{aligned}$$

This strategy profile is a Nash equilibrium, and the equilibrium outcome is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_1; i_4 - h_1; i_5 - h_2; i_6 - h_2\}.$$

Now i_4 improves her score in a new score list as $\pi_2(i_1) > \pi_2(i_2) > \pi_2(i_4) > \pi_2(i_3) > \pi_2(i_5) > \pi_2(i_6)$ and privilege scores $\pi_2(i_1) > \pi_2(i_2) > \pi_2(i_4) + \bar{s} > \pi_2(i_3) + \bar{s} > \pi_2(i_5) > \pi_2(i_6)$. Then students report preferences as follow:

$$\begin{aligned} i_1, i_2, i_3, i_4 :&\quad \text{ES:}\quad h_1 P_i \emptyset \cdots | \quad \text{NA: } \{h_1 | h_1 P_i h_2, \cdots\},\\ i_5, i_6 :&\quad \text{ES:}\quad \cdots | \quad \text{NA: } \{h_1 P_i h_2\}. \end{aligned}$$

This strategy profile is also a Nash equilibrium, and the equilibrium outcome is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_2; i_4 - h_2; i_5 - h_1; i_6 - h_1\}.$$

In this example, i_4 is strictly worse off although her standing is improved in π_2.

Example 3 shows that the concerted action between i_1 and i_2 may result in another student receives a strict worse off matching result, even this student's standing is improved. In reality, since the incentive for a middle school to achieve a better collective matching outcome cannot be ruled out in the CESM, more cities begin to switch to a new mechanism in which the early selection is abandoned and the privileges are assigned to students only in the normal admission stage. Next we begin to address this type of mechanisms.

2.2 Chinese Type-Specific Mechanism (CTSM)

The CTSM also follows an adjusted Chinese parallel mechanism with permanent-execution vector $\mathbf{e} = (e_1, e_2, \ldots)$. There are two key differences between the CTSM and the aforementioned CESM. First, the CTSM only has one matching stage, i.e. middles schools don't select indexed students before the normal admission. Second, students have no restriction to choose their ROLs in the matching algorithm, and the top q^p_{hm} students who choose school h as their first choices from a middle school m automatically receive the privilege score $\bar{s}$ in addition to their exam scores π. More precisely, the algorithm selects the matching outcomes as follow:

Round 1:

- Then each student applies to her first choice. For each high school h, it groups applicants based on their middle schools. Then for those from the same middle school m, the top q^p_{hm} applicants according to their exam scores receive a lump-sum privilege score $\bar{s}$, i.e. their priority scores become $\pi(i) + \bar{s}$; and other students' priority scores are still π. After that, school h pools all applicants together, then holds the top q_h applicants based on their priority scores and reject other applicants. Furthermore, all rejected students' priority scores are changed into their exam scores (the priority scores are equal to π).

  In general:

- Each rejected student i who has not yet applied to her (e_1)th-choice school applies to her next-preferred school. A student who has been rejected by all her first e_1 choices does not apply to any other schools until the next round. Each school h reviews the new applicants, along with those currently held in the admission pool, and then tentatively holds the top q_h applicants in its admission pool based on the priority scores. The other applicants are rejected.

- The round terminates whenever each student either is held in a school's pool or has been rejected by all her first e_1 choices. At this point, all tentative assignments become final. For each school h, its remaining total quotas are denoted $q_{h,2}$.

In general

Round $k > 1$

- Each student applies to her $\left(\sum_{j=1}^{k-1} e_j + 1\right)$th-choice school. Then, as in Round 1, each school j tentatively holds the top $q_{h,k}$ applicants in the admission pool (again, based on priority scores).

  In general:

- Each rejected student i who has not already applied to her $\left(\sum_{j=1}^{k} e_j\right)$th-choice school applies to her next-preferred school. A student who has been rejected by all her first $\sum_{j=1}^{k} e_j$ choices does not apply to any other schools until the next round. Each school j reviews the new applicants, along with those currently held in the pool, and tentatively holds the top $q_{h,k}$ applicants in its pool based on priority scores.

The algorithm terminates when each student is admitted to a school and all the tentative assignments are final.

Similar as the CESM, we focus on the analysis of the CTSM when the DA mechanism is adopted in the matching procedure, i.e. $\mathbf{e} = (\infty, \infty, \ldots)$ (see Appendix for the general case). Although the CTSM abandons the early selection stage, it still suffers deficiencies from the variation of priority scores. In the next example, we first indicate the CTSM is not strategy-proof. Since a student enjoys the privilege of a high school only if she chooses this school as the first choices, if her favorite school is too competitive or does not provide privilege seat, she may strategically put her second favorite school, from which she may gain the privilege additional score, as the first choice to increase the chance to be admitted.

Example 4. The CTPM is not strategy-proof: There are two high schools h_1, h_2 and h_3, and two middle schools m_1 and m_2 with $\mathbf{q}_{h_1} = \{2, 0, 1\}$, $\mathbf{q}_{h_2} = \{1, 0, 1\}$ and $\mathbf{q}_{h_3} = \{1, 0, 0\}$. Students $i_1, i_2 \in m_1$ and $i_3, i_4 \in m_2$ under exam scores $\pi(i_1) > \pi(i_3) > \pi(i_2) > \pi(i_4)$ and privilege scores $\pi(i_3) + \bar{s} > \pi(i_4) + \bar{s} > \pi(i_1) > \pi(i_2)$. All students have the same preferences on schools: $h_1 P_i h_2 P_i h_3$. If students report their true preferences, then the matching outcome under the CTPM is

$$\{i_1 - h_1; i_2 - h_2; i_3 - h_1; i_4 - h_3; \}.$$

If i_4 reports $h_2 P_{i_4} h_1 P_{i_4} h_3$ instead, then the matching outcome becomes

$$\{i_1 - h_1; i_2 - h_3; i_3 - h_1; i_4 - h_2; \}.$$

In this outcome, i_4 gets into a more desired school by manipulating his preference.

Since no dominant strategy might be chosen by all students under the CTSM, hence our analysis will focus on Nash equilibrium outcomes of the CTSM. Similar as the CESM, the equilibrium outcome of the CTSM is still weakly stable.

Proposition 2. *A Nash equilibrium outcome of the CTPM is weakly stable.*

In the next two examples, we show that the CTSM neither is minimally responsive nor respect improvement.

Example 5. The CTSM is not minimally responsive. This example is a simple adjustment of Example 2. There are three high schools h_1, h_2 and h_3 and two middle schools m_1 and m_2 with $\mathbf{q}_{h_1} = \{3, 1, 0\}$, $\mathbf{q}_{h_2} = \{1, 0, 0\}$ and $\mathbf{q}_{h_3} = \{1, 0, 0\}$. Students $i_1, i_2, i_3 \in m_1$ and $i_4, i_5 \in m_2$ under exam scores $\pi(i_5) > \pi(i_1) > \pi(i_2) > \pi(i_4) > \pi(i_3)$ and privilege scores $\pi(i_5) > \pi(i_1) > \pi(i_2) > \pi(i_3) + \bar{s} > \pi(i_4)$. i_5 has the preference $h_3 P_{i_5} h_1 P_{i_5} h_2$, all other students have the same preferences $h_1 P_i h_2 P_i h_3$.

Students report their preferences as follow:

$$\begin{aligned} i_1, i_2: &\quad \{h_3 P_i h_1 P_i h_2\}, \\ i_3: &\quad \{h_1 P_i h_2 P_i h_3\}, \\ i_4: &\quad \{h_1 P_i h_2 P_i h_3\}, \\ i_5: &\quad \{h_3 P_i h_1 P_i h_2\}, \end{aligned}$$

This strategy profile is an Nash equilibrium, and the matching outcome of the CESM is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_1; i_4 - h_2; i_5 - h_3\}.$$

Now if m_1 receives one more privilege seat from h_1, i.e. $\mathbf{q}_{h_1} = \{3, 2, 0\}$, and students report their preferences as follow:

$$\begin{aligned} i_1, i_2: &\quad \{h_1 P_i h_2 P_i h_3\}, \\ i_3: &\quad \{h_1 P_i h_2 P_i h_3\}, \\ i_4: &\quad \{h_1 P_i h_2 P_i h_3\}, \\ i_5: &\quad \{h_3 P_i h_1 P_i h_2\}, \end{aligned}$$

This strategy profile is an Nash equilibrium, and the matching outcome of the CESM is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_2; i_4 - h_1; i_5 - h_3\}.$$

i_3 is strictly worse off in the second case and no other students in m_1 are better off when h_1 gives one more privilege seat to m_1.

Example 5 indicates that eliminating the early selection stage does not overcome the shortcoming of group collusion. In the first equilibrium strategy, i_1 and i_2 strategically choose their least favorite schools h_3 as the first choice and get rejected, this strategy give i_3 an opportunity to receive the privilege score for her first choice h_1, because students can receive privilege score for their first choice. In the second equilibrium, when i_1 and i_2 both choose h_1 as their first choices, i_3 will be rejected by h_1. This result implies the variation of priority with rounds causes the CTSM fails to satisfy the minimal responsiveness. Because of the same reason, we will show that the CTSM does not respect improvement in the next example.

Example 6. Equilibrium outcome of the CTSM may not be respect improvement: There are two high schools h_1 and h_2 and h_3 and two middle schools m_1 and m_2 with $\mathbf{q}_{h_1} = \{4, 2, 0\}$ and $\mathbf{q}_{h_2} = \{2, 0, 0\}$ and $\mathbf{q}_{h_3} = \{1, 0, 0\}$. Students $i_1, i_2, i_3, i_4 \in m_1$ and $i_5, i_6, i_7 \in m_2$ under exam scores $\pi_2(i_7) > \pi_2(i_1) > \pi_2(i_2) > \pi_2(i_3) > \pi_2(i_4) > \pi_2(i_5) > \pi_2(i_6)$ and privilege scores $\pi_2(i_7) > \pi_2(i_1) > \pi_2(i_2) > \pi_2(i_3) + \bar{s} > \pi_2(i_4) + \bar{s} > \pi_2(i_5) > \pi_2(i_6)$. i_7's preferences is $h_3 P_i h_1 P_i h_2$ All other students have the same true preferences on high schools, i.e. $h_1 P_i h_2 P_i h_3$. Students report preferences as follow:

$$\begin{aligned} i_1, i_2 :&\quad \{h_3 P_i h_1 P_i h_2\}, \\ i_3, i_4 :&\quad \{h_1 P_i h_2 P_i h_3\}, \\ i_5, i_6 :&\quad \{h_1 P_i h_2 P_i h_3\}, \\ i_7 :&\quad \{h_3 P_i h_1 P_i h_2\}. \end{aligned}$$

This strategy profile is a Nash equilibrium, and the equilibrium outcome is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_1; i_4 - h_1; i_5 - h_2; i_6 - h_2; i_7 - h_3\}.$$

Now i_4 improves her score in a new score list as $\pi_1(i_7) > \pi_1(i_1) > \pi_1(i_2) > \pi_1(i_4) > \pi_1(i_3) > \pi_1(i_5) > \pi_1(i_6)$ and privilege scores $\pi_1(i_7) > \pi_1(i_1) > \pi_1(i_2) > \pi_1(i_4) + \bar{s} > \pi_1(i_3) + \bar{s} >$

$\pi_1(i_5) > \pi_1(i_6)$. Then students report preferences as follow:

$$\begin{aligned} i_1, i_2, i_3, i_4 : &\quad \{h_1 P_i h_2 P_i h_3\}, \\ i_5, i_6 : &\quad \{h_1 P_i h_2 P_i h_3\}. \\ i_7 : &\quad \{h_3 P_i h_1 P_i h_2\}. \end{aligned}$$

This strategy profile is a also Nash equilibrium, and the equilibrium outcome is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_2; i_4 - h_2; i_5 - h_1; i_6 - h_1; i_7 - h_3\}.$$

In this example, i_4 is strictly worse off although her standing is improved in π_1.

We have indicated that both the CESM and the CTSM suffer deficiencies due to the dynamic procedure and/or the variation of priority. Next, we propose a simple new mechanism that keeps the type-specific privilege for the affirmative action policy and overcome the flaws in the real world mechanism.

2.3 Student Optimal Type-Privilege Mechanism (SOTPM)

The SOTPM is an adjustment of the deferred-acceptance mechanism and keeps the middle school specific privilege. In particular, the privileges are assigned to top students in each middle school regardless their ROLs. More precisely, the priority score $\hat{\pi}$ under the SOTPM is defined as follow: for student $i \in m$, $\hat{\pi}(i) = \pi(i) + \bar{s}$ if student i's exam score $\pi(i)$ is among the top q_m^p of all students in middle school m, and $\hat{\pi}(i) = \pi(i)$ otherwise. Then, the DA mechanism is used to match students to schools. The algorithm selects the matching outcomes as follow:

- **Round 1:** Then each student applies to her first choice. For each high school h, it holds the top q_h applicants based on their priority scores ($\hat{\pi}$) tentatively and reject other applicants.

  In general:

- **Round k:** Each rejected student applies to her next-preferred school. For each high school h, it reviews the new applicants, along with those currently held in the admission pool school h pool, then holds the top q_h applicants based on their priority scores tentatively and reject other applicants.

The algorithm terminates when each student is admitted to a school and all the tentative assignments are final.

In the SOTPM, the privilege assignment is predetermined according to students' exam scores, and the matching algorithm of the SOTPM is exactly the same as that of the DA mechanism. Hence this mechanism shares the same properties of the DA mechanism.

Proposition 3.

1. The student optimal type-privilege mechanism is strategy-proof, furthermore it is group strategy-proof under the priorities $\hat{\pi}$; that is reporting the true preference is the weakly dominant strategy for any subgroup of I.

2. The student optimal type-privilege mechanism is stable under the priorities $\hat{\pi}$. Moreover, its matching allocation is weakly preferred by any student to any stable allocation.

Proposition 3 states the stability and strategy-proofness of the SOTPM, it also indicates the STOPM eliminates the potential group collusion problem. In the next two example, we shows that the Nash equilibrium outcome under SOTPM may dominate that under the CESM and the CTSM.

Example 7. An Nash equilibrium outcome of CESM may be Pareto inferior to that in SOTPM: There are two high schools h_1 and h_2 and two middle schools m_1 and m_2 with $\mathbf{q}_{h_1} = \{2, 0, 1\}$ and $\mathbf{q}_{h_2} = \{1, 0, 1\}$. Students $i_1 \in m_1$ and $i_2, i_3 \in m_2$ under exam scores $\pi(i_1) > \pi(i_2) > \pi(i_3)$ and privilege scores $\pi(i_2) + \bar{s} > \pi(i_3) + \bar{s} > \pi(i_1)$. Students' preferences are $h_2 P_{i_1} h_1$, $h_1 P_{i_2} h_2$ and $h_1 P_{i_3} h_2$.

Students report their preferences under the CESM as follow:

$$\begin{aligned} i_1 &: \text{ ES: } \cdots | \quad \text{NA: } \{h_1 P_i h_2 \cdots\}, \\ i_2 &: \text{ ES: } h_1 P_i h_2 P_i \emptyset | \quad \text{NA: } \{h_1 | h_1 P_i h_2 \cdots\}, \\ i_3 &: \text{ ES: } h_1 P_i h_2 P_i \emptyset | \quad \text{NA: } \{h_1 | h_1 P_i h_2; h_2 | h_2 P_i h_1\}, \end{aligned}$$

This strategy profile is an Nash equilibrium, and the matching outcome of the CESM is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_2\}.$$

Now students report their preferences under the SOTPM as follow:

$$\begin{aligned} i_1 &: \{h_2 P_i h_1\}, \\ i_2 &: \{h_1 P_i h_2\}, \\ i_3 &: \{h_1 P_i h_2\}, \end{aligned}$$

This strategy profile is an Nash equilibrium, and the matching outcome of the SOTPM is

$$\{i_1 - h_2; i_2 - h_1; i_3 - h_1\}.$$

This NE is Pareto dominates the previous NE under the CESM.

Example 8. An Nash equilibrium outcome of CTPM may be Pareto inferior to that in SOTPM: This example is a simple adjustment of Example 8. There are two high schools h_1 and h_2 and two middle schools m_1 and m_2 with $\mathbf{q}_{h_1} = \{2, 0, 1\}$ and $\mathbf{q}_{h_2} = \{1, 0, 1\}$. Students $i_1 \in m_1$ and $i_2, i_3 \in m_2$ under exam scores $\pi(i_1) > \pi(i_2) > \pi(i_3)$ and privilege scores $\pi(i_2) + \bar{s} > \pi(i_3) + \bar{s} > \pi(i_1)$. Students' preferences are $h_2 P_{i_1} h_1$, $h_1 P_{i_1} h_2$ and $h_1 P_{i_3} h_2$.

Students report their preferences under the CTSM as follow:

$$\begin{aligned} i_1: &\quad \{h_1 P_i h_2\}, \\ i_2: &\quad \{h_1 P_i h_2\}, \\ i_3: &\quad \{h_2 P_i h_1\}, \end{aligned}$$

This strategy profile is an Nash equilibrium, and the matching outcome of the CTSM is

$$\{i_1 - h_1; i_2 - h_1; i_3 - h_2\}.$$

Now students report their preferences under the SOTPM as follow:

$$\begin{aligned} i_1: &\quad \{h_2 P_i h_1\}, \\ i_2: &\quad \{h_1 P_i h_2\}, \\ i_3: &\quad \{h_1 P_i h_2\}, \end{aligned}$$

This strategy profile is an Nash equilibrium, and the matching outcome of the SOTPM is

$$\{i_1 - h_2; i_2 - h_1; i_3 - h_1\}.$$

This NE is Pareto dominates the previous NE under the CTSM.

Since the priority scores under the STOPM have no variation and are predetermined, if a student's standing is improved because of the increase of the exam score, then no other student in her middle school can change her opportunity to whether get a privilege score by a strategic action, therefore the STOPM respects improvement.

Proposition 4. *The student optimal type-privilege mechanism respects improvement under the priorities $\hat{\pi}$.*

Next, we consider another desired property for the affirmative action policy. Kojima (2012) proves an impossible theory that any quota-based affirmative action mechanism cannot be minimal responsive. It is because when the type-specific quote reserve (or restrict) for the minority students (majority students) is increased (decreased), a rejection of a majority student may trigger a rejection chain that finally hurts at least a minority student without benefit any of other minority students. The privilege-based affirmative action in the SOTPM can avoid this problem. Because, it is not guaranteed for students who receive the privileges to be admitted by their desired schools. Hence, when the quota of privileges is increased, unless some students who receive the privilege are benefited from this change, no rejection will be triggered. Therefore the SOTPM is minimal responsive in quota.

Proposition 5. *If two problem G and $\widetilde{G}$ are the same except a middle school m has a higher privilege quota from a high school h in $\widetilde{G}$ than in G, then under the SOTPM, the matching $\widetilde{\mu}$ in $\widetilde{G}$ is not Pareto inferior to the matching μ in G.*

The minimal responsiveness can be also extended to the privilege score case. Consider a problem $\widetilde{G}$ such that $\widetilde{G}$ and G are exactly the same except that the privilege score in $\widetilde{G}$ is greater than that in G, i.e. in $\bar{s}' > \bar{s}$, then we have the following result.

Proposition 6. *Under the SOTPM, the matching $\widetilde{\mu}$ in $\widetilde{G}$ is not Pareto inferior to the matching μ in G.*

Notation: The conclusion in Proposition 6 still holds when only a subset of students in the minority group receive higher priority scores in $\widetilde{G}$ than in G, as long as the relative ranking of each student within the minority group fixed.

3 Background on the Local High School Admission

The schools in our focal city can be categorized into several types based on their educational goals after students graduate from middle school. There are general high schools that prepare students for colleges and universities in China, foreign language schools (or classes) for foreign colleges or universities, fine arts schools for the fine arts colleges in China, and vocational schools for the labor market. General high schools can also be categorized into public and private high schools.

The City Education Bureau requires that all schools, regardless of type or ownership, join the centralized admission system as it pertains to middle school graduates. In addition, each student who undergoes this admission procedure must register at the school to which she is assigned by the system. Hence no outside option is available for students who intend to continue their education in this city.[1]

At the end of March in each year, the Bureau presents an admission plan that includes the quota of students that can be allocated to each school.[2] The quota for each public high school h comprises: the total quota q_h and middle school specific quota q^p_{hm} for the early selection. In early May, each middle school runs the early selection procedure to determine the indexed student identities. In mid-May, students submit their rank-ordered lists of schools. Thereafter, all students take the centralized high school entrance exam in early June. The full mark (i.e., the highest possible score) on this the exam was 665 in 2014.[3] Once the exams are graded, students are assigned to the schools by a centralized matching mechanism. All schools adopt the same strict normal priority (exam scores) over students.

Each student can list at most three schools on her ROL. The indexed student needs to list the school that she receives the privilege from as the first choice. Finally, every student must indicate whether she will accept a random assignment in the event she is rejected by her three preferred schools.

Local public high schools play a dominant role in preparing students for college. Thus, gaining entry into a public high school is the only hope most students have for attending college in China. Yet high school education in China involves more than compulsory education, and local public high schools can accommodate fewer than half of all middle school graduates. After receiving the students' ROLs and exam scores, the Bureau determines and publishes a public high school admission threshold (hereafter simply "the threshold") based on the score distribution and total available seats. Only students whose scores are above that threshold will be considered for a seat in public high schools. The threshold is meant to guarantee that the number of qualified students does not exceed the total number of available seats in public high schools. Unmatched students who have indicated acceptance of a random assignment are then randomly assigned to public high schools that still have available seats; the rest must find their own paths either to continue their schooling or to join the labor market.

The matching mechanism used by the Bureau is a special case of the Chinese early selection mechanism with permanent execution vector $(2, 1)$ in the normal admission stage. After the early selection, this mechanism's matching algorithm lasts two rounds. The first

[1]To avoid an unacceptable assignment, a student may either forgo the admission procedure or leave the application blank. Another way to avoid an undesirable assignment is to register at—but not actually attend—the assigned school. By paying additional costs, such students can instead attend schools in other cities.

[2]The admission quotas for private and vocational schools are announced at the same time.

[3]Prior to 2012, the highest possible score was 650; after 2014, it was 780.

and second choices in students' ROLs are considered in the first round, and their third choices are considered in the second round. To distinguish with the aforementioned CESM using the DA mechanism in the normal admission stage (Section 2.1), hereafter we shall reference the mechanism used in this city as the local Chinese early selection mechanism (LCESM).

4 Data Description

4.1 Data Source and Sample Selection

The data set we use consists of two parts, administrative data and survey data. The former comprise admission records in 2014. Those records include the three choices listed on students' ROLs, exam scores, final assignments, whether a student was admitted as a normal student or as a ZX student, and each student's middle school and home address. We also have some data on school characteristics: admission quotas, tuition, and dormitory accommodations.

In the administrative record, a total of 14194 students were included in the admission records. We first exclude students who were admitted by schools with special quotas, which did not affect the normal admission procedure (3.3%). Students excluded for this reason were those admitted early or by fine arts schools as well as those on sports or art scholarships.[4] Second, we exclude students whose exam scores were below the threshold (58.9%), since they were not qualified for admission to public high school. Finally, we exclude all students whose assignment outcomes were inconsistent with official rules or home addresses are missed.[5] After these exclusions, our final sample size from the administrative data was 5375.

In early May 2014, we conducted a survey of middle school graduates that asked each student to list five high schools she might attend and to rank them based on her preferences. The surveyed students were asked explicitly to report their genuine preferences, and there was no compelling reason for them not to honor this request. Because the survey was conducted just two weeks before students submitted their ROLs, it seems unlikely that their preferences would change within that short period (see Appendix **??** for details about the survey).

[4] An early admission decision is one that is made before students submit their ROLs. A student who is admitted early is still required to take the exam and to list the pre-admitting school as her first choice. Students admitted to fine arts schools must take an additional (art) exam; their admission process is handled separately from other students.

[5] For example, a few students were assigned to schools at which the cut-off was higher than their actual exam scores.

Unlike most surveys that seek to discover students' true preferences (Budish and Cantillon 2010; Kapor et al. 2017), we did not ask them to simply rank their favorite schools. Instead, respondents were asked to rank those schools they think that they might attend based on their true preferences. Recall that the exam score is the only admission criterion, and note that the highest admission cut-off may be more than 80 points higher than the lowest cut-off. Our survey design aims to avoid instances of a low-scoring student ranking schools at which she had no chance of being admitted—although such a student could list three schools with low cut-offs in her ROL. That possibility could lead to top schools being overreported in the survey, which would complicate attempts to compare the survey responses and reported ROLs of low-scoring students. The reliability of our survey is discussed further in Section 4.3.

We surveyed 6,980 students in 2014, or about half (49.17%) of the middle school graduates in that year's admission records. After we matched these students with the final administrative data sample just described—and deleted the invalid observations (e.g., students who ranked no school or only one school in the survey)—we were left with 2611 survey observations for the subsequent analysis. Thus our survey covers 48.6% of the selected sample in 2014.

4.2 School Characteristics

In the administrative data, all nonpublic high schools were coded with a single number; we therefore treated all these schools as a whole without distinctions. Table 1 summarizes the characteristics of public high schools over the study period. A total of 13 public high schools were identified, and six special classes in 2014.[6] There is a variation in the total admission quotas. The largest school can admit 600 students; at the other extreme, a small, "special class" school admits but 40 students each year.

There are eight high schools provide middle school specific privilege quotas.[7] All these high schools, except one small high school, give privileges to every middle school.[8] On average, the privilege quota accounts for 61% of the total quota of a high school. Actually, except one small high school, all other seven high schools allocate 65% of their total quota for indexed students. On average, 15% of graduates in each middle school may have the chance to become the indexed students of some high schools with a standard deviation of 1% (detailed distributions are in the Appendix). The range of the average exam scores

[6]Special classes are designed to admit gifted students and are independently operated; they also have their own admission quotas in the matching mechanism.

[7]The special classes and other five high schools do not have the privilege quotas for indexed students

[8]There is one special middle school does not receive any middles school specific privileges from high schools

of middle schools is from 361 to 538 with a mean of 445 and a standard deviation of 43. Given a large gap among the quality of middle schools and a fair even allocation of admission privileges among these schools, the local education bureau is aimed to promote more students from low quality middle school to good high schools.

To assess the reputation of public high schools, our proxy is the college admission rates of these schools in 2014. These rates are the most popular indexes used by Chinese students and parents to measure the school reputation (or quality).[9] The average of school reputation is 50.13, with a standard deviation of 37.06.[10] Other reputation-related variables, such as teachers and facilities, are unobserved in our data set. These variables can also be absorbed into schools' fix effects, which are included in our estimate in the next section. We do not take a separate approach to estimate the schools' added-value when measuring school reputation. The reason is that, when students and their families evaluate school reputation, they seldom consider value added. Instead, they use straightforward indexes as schools' rank, college admission record, or admission cut-off. Because we seek to mimic student strategies when estimating their preferences, little can be gained by considering a complicated approach to estimate the "true quality" of schools.

4.3 Student Characteristics and Behaviors

Our analysis focuses on students who were qualified to be assigned to public high schools. Approximately 90.3% of these students, whose scores were above the threshold, received seats in public high schools in 2014. These values indicate that most students who qualified for admission to take seats in the public high schools end up going there rather than entering other types of schools.

The first panel of Table 2 reports the number of schools on students' submitted ROLs. More than 93% of the normal students submit full (three-school) lists, approximately 5% of them list two schools, and fewer than 1% of all students list only one school.[11] For the indexed students, 91% of them list three schools in the ROLs, 7% list two and 0.46% list one. These facts implies the indexed students are also cautious to choose their schools in the ROLs, although they receive a large advantage for their first choices.

The table's second panel shows the assignment results. 90% of the indexed students were

[9] We do not use college admission rates in 2012 or 2013 because of the missing information on a few public high schools. For schools with complete college admission records, the rates are stable with a fluctuation of approximately 2%3%.

[10] To scale the measurement in the estimate, we multiply the percentage grade by 100. For example, if the school reputation is 50% then we record it as 50 and not as 0.5.

[11] Schools that are listed twice in the same ROL are treated as a single school.

Table 1: School Characteristics

	#	Mean	s.d.	Max	Min
# of high schools	13				
# of high schools provide privileges	8				
Privilege quota of high schools*		247.25	104.72	390	26
Privilege/Total ratio*		0.61	0.11	0.65	0.33
Total Quota of high schools		278	197	600	40
Quality		83	11	97	66
# of high schools provide dorms	13				
# of Middle schools receive privileges	43				
# of Middle schools	44				
Privilege quota of middle Schools**		45.3	19.2	79	11
Privilege/Total ratio**		0.14	0.01	0.19	0.12
Total privilege quota	1946				
Total quota	5560				

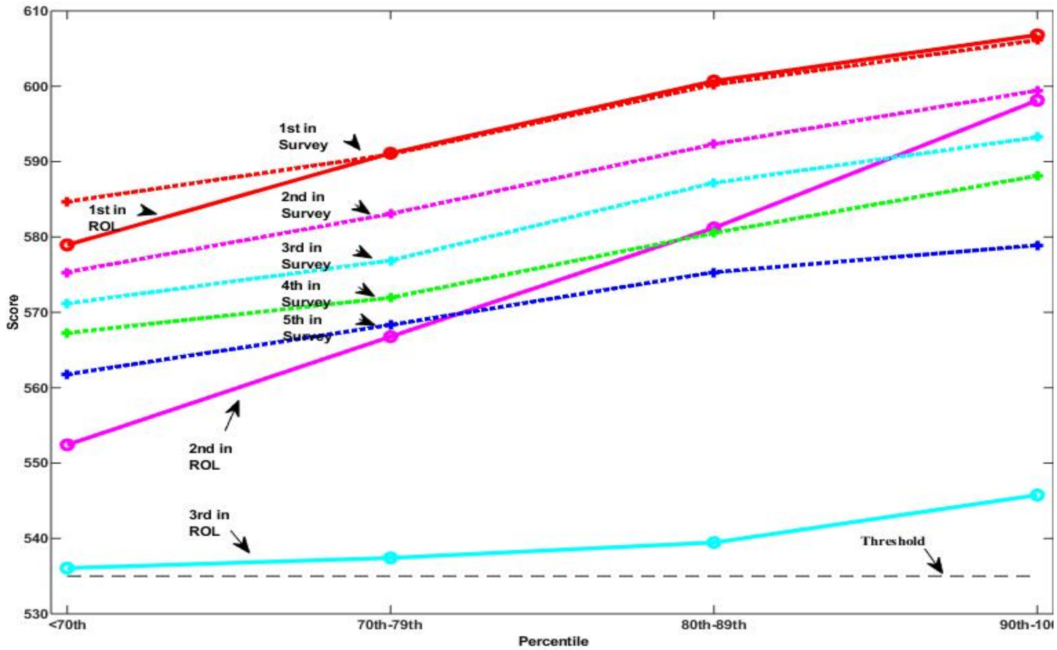
Notes * only considers the high schools providing privilege quotas. ** only considers the middles receiving privilege quotas.

assigned to their first choices, and approximately 2% and 3% of them were assigned to their second and third choices respectively, and 3% of the indexed students are rejected by all three choices. On the other hand, only 26% of the normal students were assigned to their first choices, and almost 40% of (resp. 17%) them entered their second choices (resp. third choices), and 17% were rejected by all three choices.

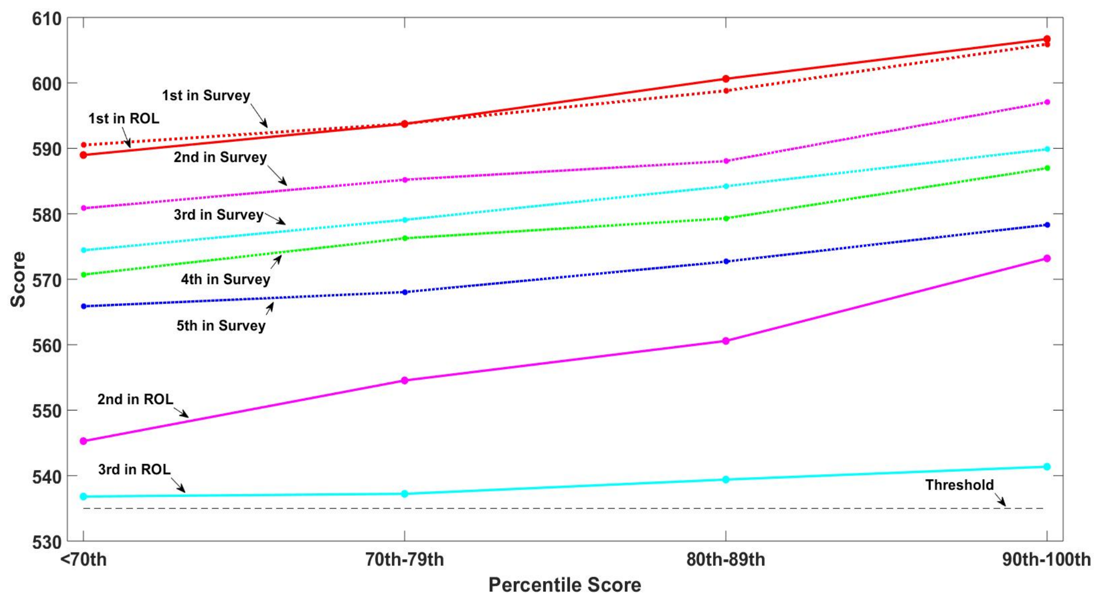
The design of the indexed students is to promote the graduates from low quality schools to get into good high schools. For students from the middle schools that have the lowest ten average exam scores 52% of these indexed students' raw exam scores are below the schools' cutoffs. In other words, these students might not get into their assigned high schools if they did not receive the privilege in the early selection stage. Relatively, this percentage is 30% for the indexed students graduating from the top ten middle schools.

Because the Chinese parallel mechanism is not strategy-proof, it is difficult to assess—while referring only to submitted ROLs—the extent to which students misrepresent their true preferences. Our survey data provide an opportunity for direct comparisons between each student's true ordinal preferences and her strategic behavior. More than 60% of the surveyed middle school graduates ranked five schools, 17% of them ranked four schools, and approximately 21% of them ranked fewer than four schools (see Table **??** in Appendix **??**).

Figure 1a shows the average admission cut-offs of schools chosen by normal students in



(a) Average cutoff for normal students



(b) Average cutoff for indexed students

Notes: The y-axis represents absolute scores, and the x-axis represents the students' exam scores in percentile. The threshold for public high school admission is 535 (60.95 percentile) in 2014.

Table 2: Student Characteristics

	Indexed Students		Normal Students	
	Freq.	Percent	Freq.	Percent
Rank Ordered Lists				
3 Schools	1789	91.93%	3100	93.71%
2 Schools	148	7.61%	189	5.71%
1 Schools	9	0.46%	19	0.57%
Assignment Results				
1st Choice	1760	90.44%	875	26.45%
2nd Choice	60	3.08%	1290	39%
3rd Choice	55	2.83%	565	17.08%
Rejected by all 3	71	3.65%	578	17.47%
Total observations	1946		3308	

the survey and the ROLs.[12] Students are grouped into four categories according to their score percentiles. In the survey, the top 10% students' exam score school cut-offs average 606.1 and 599.4 for (respectively) their first and second choices; the average cut-off for third choices (593.2) is another 6 points lower. The gaps between the third and fourth choices and the fourth and fifth choices in the survey are 5 and 9 points, respectively. The choices of students in the other three groups follow a similar pattern. Within a group, the average cut-off gap between consecutive choices is approximately 6 points and never more than 10 points. Between groups, the average cut-off for the first choice of the 80th–90th percentile students is 6 points lower than that for the highest decile of students, and this average cut-off decreases by another 9 points (to 591) for the 70th–80th percentile students. The average first-choice cut-off of students below the 70th percentile of exam scores is 585. For each additional choice, average cut-offs are similarly decreasing (at a rate of 4–10 points) in exam scores.

The decline in average cut-off of students' first choice when their scores decrease indicates that the surveyed students answered our questions truthfully by listing and ranking schools to which they might actually be admitted. The gaps between consecutive choices within groups in the survey indicate that student preferences w.r.t. schools were decreasing in the popularity of those schools; in 2014, the consecutive cut-off gaps for two popular schools were between 3 and 9 points. Also, the small cut-off gaps (4–10 points) between consecutive

[12]The corresponding table can be found in Appendix **??**.

choices within each group implies that the preferences reported in the survey are reliable enough to be viewed as the students' true preferences.

In the rank-ordered lists, the average cut-offs for the first choices of students whose exam scores were above the 70th percentile nearly coincide with the corresponding parts in the survey, although the average cut-offs for the first choices of low-scoring students (i.e., with exam scores below the 70th percentile) are 6 points lower than in the survey. However, the gap between the first and second choices increases significantly with declining exam scores. The gap in the average cut-offs between the first and second choices for the top 10% students is almost the same as that in the survey, but this gap increases to 19 points for the 80th–90th percentile students and to about 25 points for the two groups of low-scoring students. Finally, the average cut-offs for third choices are consistently close to the threshold (of 535) for all groups in the ROLs.

When compared with the survey data, the large gaps between consecutive ROL choices reveal students' strategic behavior in their submitted preferences: maintaining a sufficiently large gap between choices toward the end of increasing their chances of being admitted to some school.[13] The coincidence between the first choices in the survey and the ROLs indicates that students prefer applying to their favorite attainable schools. This coincidence, and the small cut-off gaps among choices reported in the survey, provide further evidence that the surveyed students accurately reported their five favorite attainable schools. Yet students, and especially those who were not in the top-scoring group, strategically manipulated their reported preferences in the ROLs so as to increase their overall likelihood of being admitted—that is, in the event of being rejected by their first choices. Thus the second choices in the ROLs of 80th–90th percentile (resp., 70th–80th percentile) students are close to their fourth (resp., fifth) choices in the survey. Moreover, most students (across all four groups) chose a leftover school as their third choice because the ROL is restricted to only three choices.

For the indexed students, their choices in the survey and the ROLs have similar pattern as the normal students with one obvious difference (Figure 1b). First, the average gaps between their first and second choices in the ROLs are much larger than those for normal students. This gap is 33 points for the top 10% indexed students, and it increases to around 40 points for other groups of students. This result reflect the fact that indexed students need to keep a large gap between choices because they will lose the 30 points privilege once they are rejected by their first choices.

[13]This finding is consistent with the literature that suggests students behave strategically under non–strategy-proof mechanisms (see e.g. Abdulkadiroğlu et al. 2005; Chen and Sönmez 2006; Abdulkadiroğlu et al. 2017).

5 Empirical Model and Preference Estimate

Student i's (indirect) utility from being assigned to public high school j is

$$u_{i,j} = \sum_{l} \beta^l y_j^l + \sum_{w} \beta^w x_i^w y_j^w + \beta^D f(d_{ij}, X_i, Y_j) + \varepsilon_{ij} \tag{1}$$

and that the utility from being assigned to nonpublic high school o is

$$u_{i,o} = F_o + \varepsilon_{io}. \tag{2}$$

Here $Y_j \equiv \{y_j\}$ is a vector of school j's observed characteristics; $X_i \equiv \{x_i\}$ is a vector of student i's observed characteristics; d_{ij} is the home–school distance;[14] F_o is the fixed effect of nonpublic high schools; and ε_{ij} and ε_{io} are i's idiosyncratic taste for (respectively) public high school j and nonpublic high schools. In the estimate, we assume that the home–school distance is additively separable and independent of unobserved student preferences; in addition, we normalize the coefficient d_{ij} for the home–school distance to be -1.[15]

We do not present the random coefficient model for estimating students' heterogeneous preferences for observed school characteristics (as in, e.g., Abdulkadiroğlu et al. 2015; Agarwal and Somaini 2018) owing to our data's limited variation. In China, a general high school's sole education goal is to prepare students for the college entrance exams. Except with regard to quality, schools' observed characteristics—for example, facilities—are fairly homogeneous. Even their teaching programs are fully controlled by the local education bureau. Furthermore, students who are qualified to gain seats in local public high schools exhibit similar preferences for schools (see Appendix **??** for details of students' survey responses). To avoid the mistake of choosing the wrong empirical model, we present an alternative random coefficient model in Appendix **??** and then compare the resulting estimates; the random coefficient model performs worse than does the nonrandom coefficient model on both the within-sample and the out-of-sample test.

We follow Abdulkadiroğlu et al. (2015) in not explicitly modeling an outside option. The reason for this choice is that, as mentioned in Section 3, no outside option can be observed in the current admission record. In addition, we make the following assumption.

[14] The road distance d_{ij} is calculated via Google Maps by inputting the focal school's address and the student's home address.

[15] Unlike admission to elementary and middle schools, the high school admission procedure does not consider the locations of school districts or homes. Hence we assume that, in this city, the school choice mechanism does not directly influence residential decisions or local housing prices.

Assumption 1. *The terms ε_{ij} and ε_{io} are independent of the explanatory variable, X_i, Y_j, d_{ij}, and F_o. Both ε_{ij} and ε_{io} are independent and identically distributed (i.i.d.) and exhibit a type I extreme value distribution with cumulative distribution function (CDF) $F(\varepsilon)$.*

We use the survey data to estimate student preferences. The advantage of survey data is that our estimates can proceed without having to account for students' strategic behavior when they submit their ROLs. Each surveyed student ranked five schools that she believed herself capable of attending. This procedure implies that the student first selects the schools for which admission is a distinct possibility and then, after identifying those schools, ranks them. That process complicates our constructing a model of how these middle school graduates select schools in the first place. For example, if a school with a high admission cut-off does not make the surveyed student's list, then it is difficult to distinguish between (a) her preferring the listed schools to the focal school and (b) her thinking that admission to the high–cut-off school is not possible. From the evidence presented in Section 4.3, we conclude that the survey responses reflect students' true preferences—that is, conditional on their belief in the possibility of admission. To simplify the estimation process, we focus on the listed schools' ranks in the survey (i.e., without considering the unlisted schools). In other words, we do not attempt to infer the relative ranks of listed and unlisted schools.

While referring to the survey data, we use the rank-ordered logit model (Beggs et al. 1981) to estimate coefficients.[16] Given a surveyed student i's ranked school list $(j_1, \ldots, j_{l_i})_i$ of length $l_i \leq 5$, we conclude that j_1 is her favorite school among all the l_i schools on her survey list, that j_2 is her second-favorite school, and so on. The joint probability of these choices is

$$\Pr(u_{i,j_1} > u_{i,j_1} > \cdots > u_{i,j_{l_i}}) = \prod_{k=1}^{l_i-1} \frac{e^{\mu_{i,j_k}}}{e^{\mu_{i,j_k}} + e^{\mu_{i,j_{k+1}}} + \cdots + e^{\mu_{i,j_{l_i}}}}, \tag{3}$$

where $\mu_{i,j}$ is the deterministic component of $u_{i,j}$ or $u_{i,o}$.[17] Then the log-likelihood function can be written as

$$\log L_1(\boldsymbol{\beta}) = \sum_{i=1}^{n} \sum_{k=1}^{l_j-1} \mu_{i,j_k} - \sum_{i=1}^{n} \sum_{k=1}^{l_i-1} \log \left(\sum_{s=k}^{l_i} e^{\mu_{i,j_s}} \right). \tag{4}$$

Now we can estimate coefficients by using maximum likelihood estimation.[18]

[16] Because $c_{ij} = c_0$ in this step, $\boldsymbol{\alpha}$ does not appear in the utility function.

[17] More precisely, $\mu_{i,j} = \sum_l \beta_l y_j^l + \sum_w \beta_w x_i^w y_j^w + \beta_D f(d_{ij}, Y_j)$ when j is a public high school and $\mu_{i,j} = F_o$ when j is not a public high school.

[18] We assume that the utility function has an additively separable form; it is therefore easy to show that $\log L_1$ is globally concave in the parameters—from which it follows that there exists a unique maximum of the likelihood function.

5.1 Estimation Results

Table 3 presents the estimated coefficients for the utility function. Odd and even columns report the results for normal and indexed students respectively. Columns 1 to 4 report the results when student–school interaction terms are not considered; Columns 7 and 8 provide results for the full model without the school fixed effect. We focus on Columns 5 and 6, which correspond to the full model with the school fixed effect for normal and indexed students, respectively.

Rows 2–4 of Columns 5 and 6 report student preferences regarding school quality. Students are classified into the following three groups based on their exam scores: high-scoring students, whose scores are above the 90th percentile; medium-scoring students, whose scores are between the 70th and 90th percentile; and low-scoring students, whose scores are below the 70th percentile but above the threshold. Among normal students (Column 5), the top students are much more sensitive to school quality than those in the other two groups. For example, if school quality increases by 1 unit, then, high-scoring girls displays willingness to travel an additional distance of nearly 0.54 kilometers; the corresponding distances for medium- and low-scoring girls are 0.2 km and 0.18 km, respectively. In the same situation, high-scoring (resp., medium- and low-scoring) boys are willing to travel an additional 2.75 km (resp., 1.03 and 0.92 km). For indexed students, no obviously different attitudes are reported across student groups. If school quality increases by 1 unit, girls are willing to travel an additional distance of 0.3 to 0.35 kilometers, and boys are willing to travel an additional 0.86 to 1 km. For high-scoring students, the normal students are more sensitive than their counterpart indexed students. One possible explanation is that certain normal students who may have the chance to be the indexed students of specific schools would rather be normal students to compete for desired schools; this scenario implies that such students value school quality more than high scoring indexed students who want a safe option. For medium- and low-scoring students, the indexed students value the school quality more than the same score normal students. One possible explanation is that these indexed students, who genuinely benefited from the affirmative action policy, especially from relatively low quality middle schools, have a greater desire to enter good high schools than their classmates from the similar scoring group.

A variation across groups exists for students valuation of school capacity, which we normalize to 100 seats. All students prefer small schools when other variables are fixed, but medium-scoring normal students dislike large schools the most. When school capacity decreases by 100 seats, students are willing to travel an additional 1.54 km; however, high-scoring (resp., low-scoring) normal students are willing to travel 0.94 km (resp., 1.19 km) farther. For the indexed students, medium-scoring students are willing to travel an additional 2.34 km when school capacity decreases by 100 seats, and low-scoring students are willing to travel 2.45 km.

Table 3 reports our estimates for other parameters. Rows 6–8 of Columns 5 and 6 depict that high-scoring students have a somewhat unfavorable attitude toward special classes for normal and indexed students; by contrast, such classes are viewed positively by the other two student groups. Rows 9–10 reveal that a student's utility from attending a school increases when her exam score is close to (i.e., within 15% of) the average for other students admitted there, and indexed students tend to attend a school within her score range. This outcome reflects peer pressure in schools. Rows 15–16 of Column 3 indicate that a school's provision of dormitory accommodation can reduce normal students' negative concerns about travel distance, especially for girls, but this dorm effect is opposite for the female indexed students, despite the small magnitude.

5.2 Model Fit

Next we examine how well our preference estimates match the data. We conduct the out-of-sample tests to check the aggregate-level matching patterns. Table 4 compares the actual and predicted admission cut-offs of each high school.[19]

For the out-of-sample test, Column 2 of the table reports the schools' predicted cut-offs for year 2014. With only one exceptions, the gaps between the actual and predicted cut-off are less than 1% of the full mark (665). The predicted results also correctly identify all the leftover schools, for which the cut-off is 530.

We also explore the aggregate-level matching patterns for students' first two school choices in Table 5. For our out-of-sample test, the data show that 26.45% of normal students were admitted by their first-choice schools; our predictions are, respectively, 27.77%. We underpredicted (5%) the total number of normal students who were admitted by their second choices. For the indexed students, the data show that 90.44% of them got into their first choices, and our prediction is 93.9%. We also predict that 2.8% of the indexed students are admitted by their second choices and the actual number is 3%.

[19]Reported results are the admission cut-offs for the first round. The actual second-round cut-offs of all popular schools are infinity while those of all leftover schools are equal to the threshold. Given that our predicted results correctly identify all popular and leftover schools, we report results only for the first-round cut-offs.

Table 3: Preference Parameters

	No student interactions				With student interactions			
	Normal	Indexed	Normal	Indexed	Normal	Indexed	Normal	Indexed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Quality	0.835	0.863	0.296	0.401				
	(0.011)	(0.064)	(0.020)	(0.028)				
Quality × H					0.539	0.303	0.688	0.584
					(0.155)	(0.099)	(0.032)	(0.120)
Quality × M					0.201	0.352	0.376	0.483
					(0.038)	(0.028)	(0.012)	(0.068)
Quality × L					0.181	0.323	0.361	0.319
					(0.030)	(0.063)	(0.014)	(0.075)
Special class	-1.006	4.821	-2.121	5.218				
	(0.325)	(1.841)	(1.015)	(2.052)				
Special class × H					-6.675	-2.118	-2.657	-1.699
					(1.972)	(1.731)	(0.560)	(1.072)
Special class × M					0.602	3.710	1.204	3.133
					(1.592)	(2.108)	(0.342)	(2.234)
Special class × L					6.504	8.640	5.300	16.302
					(5.591)	(5.140)	(1.193)	(2.234)
Score range					0.597	1.595	0.216	1.380
					(0.430)	(0.592)	(0.183)	(0.769)
Score range × Male					0.315	-0.386	0.898	-1.670
					(0.550)	(0.933)	(0.220)	(0.879)
Same district					-1.896	-1.520	-2.401	-2.209
					(0.247)	(0.315)	(0.107)	(0.386)
Same district × Male					1.739	2.074	2.586	3.171
					(0.309)	(0.442)	(0.143)	(0.433)
Distance	-1	-1	-1	-1	-1	-1	-1	-1
Distance × Male					0.804	0.648	0.933	0.914
					(0.034)	(0.052)	(0.010)	(0.029)
Dorm	-3.924	-4.415	4.253	0.205	4.445	-0.143	-0.907	-1.621
	(0.119)	(0.542)	(0.967)	(1.473)	(1.095)	(0.840)	(0.137)	(0.495)
Dorm × Male					0.684	1.171	0.756	1.603
					(0.307)	(0.542)	(0.164)	(0.553)
Capacity	-0.011	-1.189	-1.969	-2.554				
	(0.055)	(0.265)	(0.136)	(0.213)				
Capacity × H					-0.941	-1.311	0.217	-0.054
					(0.835)	(0.410)	(0.318)	(0.514)
Capacity × M					-1.542	-2.342	-0.632	-1.748
					(0.291)	(0.494)	(0.081)	(0.317)
Capacity × L					-1.190	-2.450	-0.540	-2.293
					(0.237)	(0.272)	(0.062)	(0.523)
Indexed High School						3.632		3.249
						(0.281)		(0.314)
Non-public high school	43.909		2.005	3.372	1.347	3.403	13.364	3.249
	(0.946)		(0.799)	(0.480)	(0.653)	(0.631)	(1.115)	(0.314)
School Fixed Effect			Y	Y	Y	Y		

Notes: Standard errors in parentheses. Distance is measured by kilometer. Both normal and ZX quotas are normalized to 100 seats. Tuition is normalized to 1000 Yuan. H, M and L represent high-scoring, medium-scoring and low-scoring students respectively.

Table 4: Admission Cutoffs

School ID	(1) True Cutoffs	(2) Predicted	(3) Diff.
141	605.0	597.8	7.2
142*	535.0	535.0	0.0
147	558.0	562.5	-4.5
167	593.5	589.8	3.7
173	552.0	554.9	-2.9
179	573.5	573.7	-0.2
181*	535.0	535.0	0.0
183	611.0	605.0	6.0
184*	535.0	535.0	0.0
185	583.0	579.6	3.4
186	576.0	577.8	-1.8
187	596.0	593.3	2.7

Notes: This table indicates the out-of-sample test for the schools' cutoffs. The full mark is 665. The threshold is 535 in 2014. * indicates the leftover schools with cutoff equal to the threshold.

Table 5: Admission Patterns (%)

	Out of Sample		
	Data 2014	Predicted	Diff.
Normal 1st Choice	26.45	27.77	-1.32
Normal 2nd Choice	39	33.48	5.52
Indexed 1st Choice	90.44	93.9	-3.46
Indexed 2nd Choice	3.08	2.8	0.28

Notes: This table indicates the out-of-sample test of the matching patterns for the 1st and 2nd choices for both normal and indexed students in the ROLs.

6 Counterfactual Analysis

Using the estimated preferences, we simulate the students' application lists. In the simulation, we use the profiles of students and schools from the 2014 administrative data. To analyze the welfare effect of different mechanism, we use the matching outcomes under the SOTPM as our benchmark. We analyze the welfare changes when the CTSM, CESM, and LCESM replace the DA mechanism. Under the SOTPM mechanism, we assume that students' ROLs report their true preferences; under other mechanisms, we create ROLs that reflect each student's best response in equilibrium (see Appendix **??** for details). We use 1,000 simulations in which each student experiences a different vector of random utility shocks.

6.1 Students' Welfare

When the SOTPM is replaced by the CTSM, the average welfare of students falls to less than 1% (Table 6). The high-scoring students experience a welfare gain by 2% on average, and medium- and low- scoring students have a welfare loss by less than 2%. Table 7 identifies the percentage of "winners" (whose welfare increase) and "losers" (whose welfare decrease) when the SOTPM is replaced. Under the CTSM, the proportion of winners is 12%, and the proportion of losers is 14%. High-scoring students have more winners (18%) than losers (8.75%); however, more losers (21.8%) than winners (13.7%) are found among medium-scoring students. For low-scoring students, the number of winners is equal to that of losers (3%), which implies that the CTSM benefits high-scoring students rather than medium-scoring students.

When the SOTPM mechanism is replaced with the CESM, the changes in student welfare are different with the case in the CTSM. Overall, students have a welfare gain by 7.8%.

However, the high- and medium-scoring students experience a welfare loss, especially, the medium-scoring students surface 43% welfare loss on average. On the contrary, the low-scoring students benefit through this mechanism by 136%. More precisely, there are 27% losers (resp., 70%) relative to 5.7% (resp., 24.5%) winners among high-scoring students (resp., low-scoring students) under the CESM, while 94% of low-scoring students become winners and only 0.2% of them become losers.

When the SOTPM is replaced with the LCESM, the average welfare of students falls by 2%. Similar as the CESM case, the high- and medium-scoring students experience a welfare loss on average. High-scoring students on average lose the welfare by 1%, and 11% of them become winners and 26% of them suffer a welfare loss. The medium-scoring students suffer a welfare loss by 18% on average with 11% winners and 63% losers. The low-scoring groups continue to benefit through the LCESM with a 33% welfare gain, nearly 50% of them are winners and 37% are losers.

Table 6: Change of Welfare (%)

	CTSM	CESM	LCESM
Overall	-0.79	7.82	-2.02
H-scoring	2.02	-6.10	-1.19
M-scoring	-1.84	-43.41	-18.18
L-scoring	-1.52	136.60	32.96

Notes:

Table 7: Winners and Losers (%)

	CTSM		CESM		LCESM	
	winner	loser	winner	loser	winner	loser
Overall	12.37	14.19	35.94	43.39	20.13	47.42
H-scoring	18.12	8.75	5.76	27.17	11.17	26.11
M-scoring	13.73	21.80	24.49	70.75	11.25	62.39
L-scoring	3.14	3.15	93.96	0.17	49.52	37.20

Notes:

Introducing the indexed student policy is aimed at providing students from low quality middle schools additional opportunities to enter into high quality high schools. Next we examine the performance of different mechanisms to achieve this goal. First, we classify the middle schools into three groups: (i) Top middle schools consist of ten middle schools with

Table 8: Diversity (%)

	Top two high schools (141+183)					Other good-quality high schools				
	DA	SOTPM	CTSM	CESM	LCESM	DA	SOTPM	CTSM	CESM	LCESM
Top middle schools	58.55	39.70	49.59	51.30	45.65	41.43	47.22	41.44	39.82	38.53
Median middle schools	36.17	51.16	41.21	40.47	43.77	51.78	42.37	48.90	52.09	49.88
Low middle schools	5.28	9.14	9.19	8.23	10.58	6.79	10.42	9.66	8.09	11.60

Notes:

the highest average entrance exam scores; (ii) low middle schools include ten middle schools with the lowest average entrance exam scores; (iii) the other middle schools are categorized as the median middle schools. In the counterfactual analysis, we consider the matching results under the DA mechanism as the benchmark.

Table 8 indicates the effects of the indexed student policy. If no such policy exists, then the students graduating from top ten middle schools take 58% of the seats in the best two high schools, and this number drops to 39.7% when the SOTPM is adopted. When other mechanisms are used, the top ten middle school graduates take from 45% to 51% of the seats in the best two high schools. Moreover, median middle school graduates contribute 36% of the seats in the best two high schools under the DA mechanism, and this number increases to 51% under the SOTPM and around 41% for the other three mechanisms. Furthermore, the SOTPM can promote 9% of the students from the low middle schools into the best two high schools from 5% under the DA mechanism. When other mechanisms are adopted, then at least 8% of low middle school graduates are admitted by the best two high schools. In brief, the indexed student policy considerably promotes students from median - and low-middle schools into the best high schools, especially the SOTPM.

For other good-quality high schools, 41% of top middle school graduates can get into these schools. This number is not changed considerably under the CTSM, CESM, or LCESM, but it is increased to 47% under the SOTPM. Students from median-middle schools, them take 51% of seats in these good-quality schools under the DA mechanism; whereas the CTSM, CESM and LCESM do not change this number more than 2%. However, the SOTPM decreases the number of admitted students from median-middle schools to 42%. For students who graduate from the low-middle schools, only 6.8% can get into these good quality high schools, and this number is promoted to 10% under the SOTPM. However, other mechanisms also increase the chances of the low-middle school graduates to enter these good quality high schools by at least 8%.

7 Conclusion

References

Abdulkadiroglu, Atila and Tayfun Sönmez (2003). “School choice: A mechanism design approach”. *The American Economic Review* 93.3, pp. 729–747.

Abdulkadiroğlu, Atila, Parag A Pathak, and Alvin E Roth (2005). “The new york city high school match”. *American Economic Review*, pp. 364–367.

Abdulkadiroğlu, Atila, Nikhil Agarwal, and Parag A Pathak (2015). *The welfare effects of coordinated assignment: Evidence from the NYC High School match*. Tech. rep. National Bureau of Economic Research.

Abdulkadiroğlu, Atila, Yeon-Koo Che, Parag A Pathak, Alvin E Roth, and Olivier Tercieux (2017). *Minimizing Justified Envy in School Choice: The Design of New Orleans' OneApp*. Tech. rep. National Bureau of Economic Research.

Agarwal, Nikhil and Paulo Somaini (2018). *Demand Analysis using Strategic Reports: An application to a school choice mechanism*. Tech. rep. 2, pp. 391–444.

Beggs, Steven, Scott Cardell, and Jerry Hausman (1981). “Assessing the potential demand for electric cars”. *Journal of econometrics* 17.1, pp. 1–19.

Budish, Eric B and Estelle Cantillon (2010). “The multi-unit assignment problem: Theory and evidence from course allocation at harvard”.

Chen, Yan and Onur Kesten (2017). “Chinese college admissions and school choice reforms: A theoretical analysis”. *Journal of Political Economy* 125.1, pp. 000–000.

Chen, Yan and Tayfun Sönmez (2006). “School choice: an experimental study”. *Journal of Economic theory* 127.1, pp. 202–231.

Doğan, Battal (2016). “Responsive affirmative action in school choice”. *Journal of Economic Theory* 165, pp. 69–105.

Doğan, Battal and Bettina Klaus (2018). “Object allocation via immediate-acceptance: Characterizations and an affirmative action application”. *Journal of Mathematical Economics* 79, pp. 140–156.

Hafalir, Isa E, M Bumin Yenmez, and Muhammed A Yildirim (2013). “Effective affirmative action in school choice”. *Theoretical Economics* 8.2, pp. 325–363.

Kapor, Adam, Christopher A Neilson, and Seth D Zimmerman (2017). “Heterogeneous Beliefs and School Choice Mechanisms”.

Kojima, Fuhito (2012). “School choice: Impossibilities for affirmative action”. *Games and Economic Behavior* 75.2, pp. 685–693.

Appendices